

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10 | 510, 508

Source: PCT

Date Processed by STIC: 10 | 16 | 04

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 10/16/2004

PATENT APPLICATION: US/10/510,508

TIME: 09:07:20

Input Set : A:\2543-1-038PCTUS - Seq List.txt

Output Set: N:\CRF4\10162004\J510508.raw

```

3 <110> APPLICANT: Terrett, Jonathan A
5 <120> TITLE OF INVENTION: Protein Involved in Cancer
7 <130> FILE REFERENCE: 2543-1-038PCT/US
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/510,508
C--> 10 <141> CURRENT FILING DATE: 2004-10-07
12 <150> PRIOR APPLICATION NUMBER: GB0208332.7
13 <151> PRIOR FILING DATE: 2002-04-11
15 <150> PRIOR APPLICATION NUMBER: GB0229875.0
16 <151> PRIOR FILING DATE: 2002-12-21
18 <160> NUMBER OF SEQ ID NOS: 6
20 <170> SOFTWARE: PatentIn version 3.1
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 1212
24 <212> TYPE: PRT
25 <213> ORGANISM: Homo sapiens
27 <400> SEQUENCE: 1
29 Met Glu Pro Arg Pro Thr Ala Pro Ser Ser Gly Ala Pro Gly Leu Ala
30 1 5 10 15
32 Gly Val Gly Glu Thr Pro Ser Ala Ala Ala Leu Ala Ala Ala Arg Val
33 20 25 30
35 Glu Leu Pro Gly Thr Ala Val Pro Ser Val Pro Glu Asp Ala Ala Pro
36 35 40 45
38 Ala Ser Arg Asp Gly Gly Gly Val Arg Asp Glu Gly Pro Ala Ala Ala
39 50 55 60
41 Gly Asp Gly Leu Gly Arg Pro Leu Gly Pro Thr Pro Ser Gln Ser Arg
42 65 70 75 80
44 Phe Gln Val Asp Leu Val Ser Glu Asn Ala Gly Arg Ala Ala Ala Ala
45 85 90 95
47 Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Ala Gly Ala Gly
48 100 105 110
50 Ala Lys Gln Thr Pro Ala Asp Gly Glu Ala Ser Gly Glu Ser Glu Pro
51 115 120 125
53 Ala Lys Gly Ser Glu Glu Ala Lys Gly Arg Phe Arg Val Asn Phe Val
54 130 135 140
56 Asp Pro Ala Ala Ser Ser Ser Ala Glu Asp Ser Leu Ser Asp Ala Ala
57 145 150 155 160
59 Gly Val Gly Val Asp Gly Pro Asn Val Ser Phe Gln Asn Gly Gly Asp
60 165 170 175
62 Thr Val Leu Ser Glu Gly Ser Ser Leu His Ser Gly Gly Gly Gly Gly
63 180 185 190
65 Ser Gly His His Gln His Tyr Tyr Tyr Asp Thr His Thr Asn Thr Tyr
66 195 200 205
68 Tyr Leu Arg Thr Phe Gly His Asn Thr Met Asp Ala Val Pro Arg Ile

```

## RAW SEQUENCE LISTING

DATE: 10/16/2004

PATENT APPLICATION: US/10/510,508

TIME: 09:07:20

Input Set : A:\2543-1-038PCTUS - Seq List.txt

Output Set: N:\CRF4\10162004\J510508.raw

```

69      210      215      220
71 Asp His Tyr Arg His Thr Ala Ala Gln Leu Gly Glu Lys Leu Leu Arg
72 225      230      235      240
74 Pro Ser Leu Ala Glu Leu His Asp Glu Leu Glu Lys Glu Pro Phe Glu
75      245      250      255
77 Asp Gly Phe Ala Asn Gly Glu Glu Ser Thr Pro Thr Arg Asp Ala Val
78      260      265      270
80 Val Thr Tyr Thr Ala Glu Ser Lys Gly Val Val Lys Phe Gly Trp Ile
81      275      280      285
83 Lys Gly Val Leu Val Arg Cys Met Leu Asn Ile Trp Gly Val Met Leu
84      290      295      300
86 Phe Ile Arg Leu Ser Trp Ile Val Gly Gln Ala Gly Ile Gly Leu Ser
87 305      310      315      320
89 Val Leu Val Ile Met Met Ala Thr Val Val Thr Thr Ile Thr Gly Leu
90      325      330      335
92 Ser Thr Ser Ala Ile Ala Thr Asn Gly Phe Val Arg Gly Gly Gly Ala
93      340      345      350
95 Tyr Tyr Leu Ile Ser Arg Ser Leu Gly Pro Glu Phe Gly Gly Ala Ile
96      355      360      365
98 Gly Leu Ile Phe Ala Phe Ala Asn Ala Val Ala Val Ala Met Tyr Val
99      370      375      380
101 Val Gly Phe Ala Glu Thr Val Val Glu Leu Leu Lys Glu His Ser Ile
102 385      390      395      400
104 Leu Met Ile Asp Glu Ile Asn Asp Ile Arg Ile Ile Gly Ala Ile Thr
105      405      410      415
107 Val Val Ile Leu Leu Gly Ile Ser Val Ala Gly Met Glu Trp Glu Ala
108      420      425      430
110 Lys Ala Gln Ile Val Leu Leu Val Ile Leu Leu Leu Ala Ile Gly Asp
111      435      440      445
113 Phe Val Ile Gly Thr Phe Ile Pro Leu Glu Ser Lys Lys Pro Lys Gly
114      450      455      460
116 Phe Phe Gly Tyr Lys Ser Glu Ile Phe Asn Glu Asn Phe Gly Pro Asp
117 465      470      475      480
119 Phe Arg Glu Glu Glu Thr Phe Phe Ser Val Phe Ala Ile Phe Phe Pro
120      485      490      495
122 Ala Ala Thr Gly Ile Leu Ala Gly Ala Asn Ile Ser Gly Asp Leu Ala
123      500      505      510
125 Asp Pro Gln Ser Ala Ile Pro Lys Gly Thr Leu Leu Ala Ile Leu Ile
126      515      520      525
128 Thr Thr Leu Val Tyr Val Gly Ile Ala Val Ser Val Gly Ser Cys Val
129      530      535      540
131 Val Arg Asp Ala Thr Gly Asn Val Asn Asp Thr Ile Val Thr Glu Leu
132 545      550      555      560
134 Thr Asn Cys Thr Ser Ala Ala Cys Lys Leu Asn Phe Asp Phe Ser Ser
135      565      570      575
137 Cys Glu Ser Ser Pro Cys Ser Tyr Gly Leu Met Asn Asn Phe Gln Val
138      580      585      590
140 Met Ser Met Val Ser Gly Phe Thr Pro Leu Ile Ser Ala Gly Ile Phe
141      595      600      605

```

## RAW SEQUENCE LISTING

DATE: 10/16/2004

PATENT APPLICATION: US/10/510,508

TIME: 09:07:20

Input Set : A:\2543-1-038PCTUS - Seq List.txt

Output Set: N:\CRF4\10162004\J510508.raw

```

143 Ser Ala Thr Leu Ser Ser Ala Leu Ala Ser Leu Val Ser Ala Pro Lys
144      610                      615                      620
146 Ile Phe Gln Ala Leu Cys Lys Asp Asn Ile Tyr Pro Ala Phe Gln Met
147 625                      630                      635                      640
149 Phe Ala Lys Gly Tyr Gly Lys Asn Asn Glu Pro Leu Arg Gly Tyr Ile
150      645                      650                      655
152 Leu Thr Phe Leu Ile Ala Leu Gly Phe Ile Leu Ile Ala Glu Leu Asn
153      660                      665                      670
155 Val Ile Ala Pro Ile Ile Ser Asn Phe Phe Leu Ala Ser Tyr Ala Leu
156      675                      680                      685
158 Ile Asn Phe Ser Val Phe His Ala Ser Leu Ala Lys Ser Pro Gly Trp
159      690                      695                      700
161 Arg Pro Ala Phe Lys Tyr Tyr Asn Met Trp Ile Ser Leu Leu Gly Ala
162 705                      710                      715                      720
164 Ile Leu Cys Cys Ile Val Met Phe Val Ile Asn Trp Trp Ala Ala Leu
165      725                      730                      735
167 Leu Thr Tyr Val Ile Val Leu Gly Leu Tyr Ile Tyr Val Thr Tyr Lys
168      740                      745                      750
170 Lys Pro Asp Val Asn Trp Gly Ser Ser Thr Gln Ala Leu Thr Tyr Leu
171      755                      760                      765
173 Asn Ala Leu Gln His Ser Ile Arg Leu Ser Gly Val Glu Asp His Val
174      770                      775                      780
176 Lys Asn Phe Arg Pro Gln Cys Leu Val Met Thr Gly Ala Pro Asn Ser
177 785                      790                      795                      800
179 Arg Pro Ala Leu Leu His Leu Val His Asp Phe Thr Lys Asn Val Gly
180      805                      810                      815
182 Leu Met Ile Cys Gly His Val His Met Gly Pro Arg Arg Gln Ala Met
183      820                      825                      830
185 Lys Glu Met Ser Ile Asp Gln Ala Lys Tyr Gln Arg Trp Leu Ile Lys
186      835                      840                      845
188 Asn Lys Met Lys Ala Phe Tyr Ala Pro Val His Ala Asp Asp Leu Arg
189      850                      855                      860
191 Glu Gly Ala Gln Tyr Leu Met Gln Ala Ala Gly Leu Gly Arg Met Lys
192 865                      870                      875                      880
194 Pro Asn Thr Leu Val Leu Gly Phe Lys Lys Asp Trp Leu Gln Ala Asp
195      885                      890                      895
197 Met Arg Asp Val Asp Met Tyr Ile Asn Leu Phe His Asp Ala Phe Asp
198      900                      905                      910
200 Ile Gln Tyr Gly Val Val Val Ile Arg Leu Lys Glu Gly Leu Asp Ile
201      915                      920                      925
203 Ser His Leu Gln Gly Gln Glu Glu Leu Leu Ser Ser Gln Glu Lys Ser
204      930                      935                      940
206 Pro Gly Thr Lys Asp Val Val Val Ser Val Glu Tyr Ser Lys Lys Ser
207 945                      950                      955                      960
209 Asp Leu Asp Thr Ser Lys Pro Leu Ser Glu Lys Pro Ile Thr His Lys
210      965                      970                      975
212 Val Glu Glu Glu Asp Gly Lys Thr Ala Thr Gln Pro Leu Leu Lys Lys
213      980                      985                      990
215 Glu Ser Lys Gly Pro Ile Val Pro Leu Asn Val Ala Asp Gln Lys Leu

```

## RAW SEQUENCE LISTING

DATE: 10/16/2004

PATENT APPLICATION: US/10/510,508

TIME: 09:07:20

Input Set : A:\2543-1-038PCTUS - Seq List.txt

Output Set: N:\CRF4\10162004\J510508.raw

216	995	1000	1005
218	Leu Glu	Ala Ser Thr Gln Phe	Gln Lys Lys Gln Gly Lys Asn Thr
219	1010	1015	1020
221	Ile Asp	Val Trp Trp Leu Phe	Asp Asp Gly Gly Leu Thr Leu Leu
222	1025	1030	1035
224	Ile Pro	Tyr Leu Leu Thr Thr	Lys Lys Lys Trp Lys Asp Cys Lys
225	1040	1045	1050
227	Ile Arg	Val Phe Ile Gly Gly	Lys Ile Asn Arg Ile Asp His Asp
228	1055	1060	1065
230	Arg Arg	Ala Met Ala Thr Leu	Leu Ser Lys Phe Arg Ile Asp Phe
231	1070	1075	1080
233	Ser Asp	Ile Met Val Leu Gly	Asp Ile Asn Thr Lys Pro Lys Lys
234	1085	1090	1095
236	Glu Asn	Ile Ile Ala Phe Glu	Glu Ile Ile Glu Pro Tyr Arg Leu
237	1100	1105	1110
239	His Glu	Asp Asp Lys Glu Gln	Asp Ile Ala Asp Lys Met Lys Glu
240	1115	1120	1125
242	Asp Glu	Pro Trp Arg Ile Thr	Asp Asn Glu Leu Glu Leu Tyr Lys
243	1130	1135	1140
245	Thr Lys	Thr Tyr Arg Gln Ile	Arg Leu Asn Glu Leu Leu Lys Glu
246	1145	1150	1155
248	His Ser	Ser Thr Ala Asn Ile	Ile Val Met Ser Leu Pro Val Ala
249	1160	1165	1170
251	Arg Lys	Gly Ala Val Ser Ser	Ala Leu Tyr Met Ala Trp Leu Glu
252	1175	1180	1185
254	Ala Leu	Ser Lys Asp Leu Pro	Pro Ile Leu Leu Val Arg Gly Asn
255	1190	1195	1200
257	His Gln	Ser Val Leu Thr Phe	Tyr Ser
258	1205	1210	

260 &lt;210&gt; SEQ ID NO: 2

261 &lt;211&gt; LENGTH: 4098

262 &lt;212&gt; TYPE: DNA

263 &lt;213&gt; ORGANISM: Homo sapiens

W--&gt; 264 &lt;400&gt; SEQUENCE: 2

265	ggtggcctct	gtggccgtcc	aggctagcgg	cggcccgcag	gcggcgggga	gaaagactct	60
267	ctcacctggt	cttgcggtcg	tggccaccgc	cggccagggg	tgtggagggc	gtgctgccgg	120
269	agacgtccgc	cgggctctgc	agttccgcgc	ggggtcgggc	agctatggag	ccgcggccca	180
271	cggcgccctc	ctccggcgcc	ccgggactgg	ccggggtcgg	ggagacgccc	tcagccgctg	240
273	cgctggccgc	agccagggcg	gaactgccc	gcacggctgt	gccctcggtg	ccggaggatg	300
275	ctgcgcccgc	gagccgggac	ggcggcgggg	tccgcgatga	gggccccgcg	gcggccgggg	360
277	acgggctggg	cagacccttg	gggcccaccc	cgagccagag	ccgtttccag	gtggacctgg	420
279	tttccgagaa	cgccgggcgg	gcccgtgctg	cggcggcggc	ggcggcggcg	gcagcggcgg	480
281	cggctggtgc	tggggcgggg	gccaagcaga	ccccgcgga	cggggaagcc	agcggcgaga	540
283	gcgagccagc	taaaggcagc	gaggaagcca	agggccgctt	ccgcgtgaac	ttcgtggacc	600
285	cagctgcctc	ctcgtcggtc	gaagacagcc	tgctcagatg	tgccggggtc	ggagtcgacg	660
287	ggcccaacgt	gagcttccag	aacggcgggg	acacgggtgt	gagcgagggc	agcagcctgc	720
289	actccggcgg	cggcggcggc	agtgggcacc	accagcacta	ctattatgat	acccacacca	780
291	acacctacta	cctgcgcacc	ttcggccaca	acaccatgga	cgctgtgccc	aggatcgatc	840
293	actaccggca	cacagccgcg	cagctggggc	agaagctgct	ccggcctagc	ctggcgggagc	900

## RAW SEQUENCE LISTING

DATE: 10/16/2004

PATENT APPLICATION: US/10/510,508

TIME: 09:07:20

Input Set : A:\2543-1-038PCTUS - Seq List.txt

Output Set: N:\CRF4\10162004\J510508.raw

295	tccacgacga	gctggaaaag	gaaccttttg	aggatggctt	tgcaaatggg	gaagaaagta	960
297	ctccaaccag	agatgctgtg	gtcacgtata	ctgcagaaag	taaaggagtc	gtgaagtttg	1020
299	gctggatcaa	gggtgtatta	gtacgttgta	tgtaaacaat	ttgggggtgtg	atgcttttca	1080
301	ttagattgtc	atggattgtg	ggtcaagctg	gaataggtct	atcagtcctt	gtaataatga	1140
303	tggccactgt	tgtgacaact	atcacaggat	tgtctacttc	agcaatagca	actaatggat	1200
305	ttgtaagagg	aggaggagca	tattatttta	tatctagaag	tctagggcca	gaatttggtg	1260
307	gtgcaattgg	tctaactctt	gcctttgcca	acgctgttgc	agttgctatg	tatgtggttg	1320
309	gattttgcaga	aaccgtgggtg	gagttgctta	aggaacattc	catacttatg	atagatgaaa	1380
311	tcaatgatat	cgaattatt	ggagccatta	cagtcgtgat	tcttttaggt	atctcagtag	1440
313	ctggaatgga	gtgggaagca	aaagctcaga	ttgttctttt	ggtgatccta	cttcttgcta	1500
315	ttggtgattt	cgtcatagga	acatttatcc	cactggagag	caagaagcca	aaagggtttt	1560
317	ttggttataa	atctgaaata	tttaatgaga	actttgggcc	cgattttcga	gaggaagaga	1620
319	ctttcttttc	tgtatttgcc	atcttttttc	ctgctgcaac	tggtattctg	gctggagcaa	1680
321	atatctcagg	tgatcttgca	gatcctcagt	cagccatacc	caaaggaaca	ctcctagcca	1740
323	ttttaattac	tacattgggt	tacgtaggaa	ttgcagtatc	tgtaggttct	tgtgttgttc	1800
325	gagatgccac	tggaaacggt	aatgacacta	tcgtaacaga	gctaacaaac	tgtacttctg	1860
327	cagcctgcaa	attaaacttt	gatttttcat	cttgtgaaag	cagtccttgt	tcctatggcc	1920
329	taatgaacaa	cttccaggta	atgagtatgg	tgtcaggatt	tacaccacta	atttctgcag	1980
331	gtatattttc	agccactctt	tcttcagcat	tagcatccct	agtgagtget	cccaaatat	2040
333	ttcaggctct	atgtaaggac	aacatctacc	cagctttcca	gatgtttgct	aaagggtatg	2100
335	ggaaaaataa	tgaacctctt	cgtggctaca	ctttaacatt	cttaattgca	cttggattca	2160
337	tcttaattgc	tgaactgaat	gttattgcac	caattatctc	aaacttcttc	cttgcacat	2220
339	atgcattgat	caatttttca	gtattccatg	catcacttgc	aaaatctcca	ggatggcgctc	2280
341	ctgcattcaa	atactacaac	atgtggatat	cacttcttgg	agcaattctt	tgttgcatag	2340
343	taatgttcgt	cattaactgg	tgggctgcat	tgctaacata	tgtgatagtc	cttgggctgt	2400
345	atatttatgt	tacctacaaa	aaaccagatg	tgaattgggg	atcctctaca	caagccctga	2460
347	cttacctgaa	tgcactgcag	cattcaattc	gtctttctgg	agtgggaagac	cacgtgaaaa	2520
349	actttaggcc	acagtgtctt	gttatgacag	gtgctccaaa	ctcacgtcca	gctttacttc	2580
351	atcttgttca	tgatttcaca	aaaaatggtg	gtttgatgat	ctgtggccat	gtacatatgg	2640
353	gtcctogaag	acaagccatg	aaagagatgt	ccatcgatca	agccaaatat	cagcgatggc	2700
355	ttattaagaa	caaatgaag	gcattttatg	ctccagtaca	tgcagatgac	ttgagagaag	2760
357	gtgcacagta	tttgatgcag	gctgctgggtc	ttggctcgat	gaagccaaac	acacttgccc	2820
359	ttggatttaa	gaaagattgg	ttgcaagcag	atatgagggg	tgtggatatg	tatataaact	2880
361	tatttcatga	tgtttttgac	atacaatatg	gagtagtggt	tattcgcta	aaagaaggtc	2940
363	tggatatatc	tcatcttcaa	ggacaagaag	aattattgtc	atcacaagag	aatctcctg	3000
365	gcaccaagga	tgtggtagta	agtgtggaat	atagtaaaaa	gtccgattta	gatacttcca	3060
367	aaccactcag	tgaaaaaaca	attacacaca	aagttgagga	agaggatggc	aagactgcaa	3120
369	ctcaaccact	gttgaaaaaa	gaatccaaag	gccctattgt	gcctttaaat	gtagctgacc	3180
371	aaaagcttct	tgaagctagt	acacagtttc	agaaaaaaca	aggaaagaat	actattgatg	3240
373	tctgggtggc	ttttgatgat	ggaggtttga	ccttattgat	accttacctt	ctgacgacca	3300
375	agaaaaaatg	gaaagactgt	aagatcagag	tattcattgg	tggaaagata	aacagaatag	3360
377	accatgaccg	gagagcgatg	gctactttgc	ttagcaagtt	ccggatagac	ttttctgata	3420
379	tcatggttct	aggagatatc	aataccaaac	caaagaaaga	aatattata	gcttttgagg	3480
381	aatcattga	gccatacaga	cttcatgaag	atgataaaga	gcaagatatt	gcagataaaa	3540
383	tgaagaaga	tgaacatgg	cgaataacag	ataatgagct	tgaactttat	aagaccaaga	3600
385	cataccggca	gatcagggtta	aatgagttat	taaaggaaca	ttcaagcaca	gctaataatta	3660
387	ttgtcatgag	tctcccagtt	gcacgaaaag	gtgctgtgtc	tagtgctctc	tacatggcat	3720
389	ggttagaagc	tctatctaag	gacctaccac	caatcctcct	agttcgtggg	aatcatcaga	3780
391	gtgtccttac	cttctattca	taaatgttct	atacagtggg	cagccctcca	gaatgggtact	3840

## VERIFICATION SUMMARY

DATE: 10/16/2004

PATENT APPLICATION: US/10/510,508

TIME: 09:07:21

Input Set : A:\2543-1-038PCTUS - Seq List.txt

Output Set: N:\CRF4\10162004\J510508.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:264 M:283 W: Missing Blank Line separator, <400> field identifier  
L:407 M:283 W: Missing Blank Line separator, <220> field identifier  
L:409 M:283 W: Missing Blank Line separator, <400> field identifier  
L:416 M:283 W: Missing Blank Line separator, <220> field identifier  
L:418 M:283 W: Missing Blank Line separator, <400> field identifier  
L:425 M:283 W: Missing Blank Line separator, <220> field identifier  
L:427 M:283 W: Missing Blank Line separator, <400> field identifier  
L:436 M:283 W: Missing Blank Line separator, <220> field identifier  
L:438 M:283 W: Missing Blank Line separator, <400> field identifier